United States Government Accountability Office

Report to the Chairman,
Committee on Health, Education,
Labor, and Pensions,
U.S. Senate

November 2014

EDUCATION AND WORKFORCE DATA

Challenges in Matching Student and Worker Information Raise Concerns about Longitudinal Data Systems

GAO-15-27

November 2014

EDUCATION AND WORKFORCE DATA

Challenges in Matching Student and Worker Information Raise Concerns about Longitudinal Data Systems

GAO Highlights

Highlights of GAO-15-27, a report to the Chairman, Committee on Health, Education, Labor, and Pensions, U.S. Senate

Why GAO Did This Study

From fiscal years 2006 through 2013, the Departments of Education and Labor provided over $640 million in grants to states through the SLDS and WDQI grant programs. These grants support states' efforts to create longitudinal data systems that follow individuals through their education and into the workforce. Analyzing data in these systems may help states improve outcomes for students and workers.

GAO was asked to review the status of grantees' longitudinal data systems. This report examines (1) the extent to which SLDS and WDQI grantees match individual student and worker records and share data between the education and workforce sectors and (2) how grantees are using longitudinal data to help improve education and workforce outcomes. To answer these questions, GAO analyzed data from a 2013 survey conducted by the DQC. This survey collected information from states on data linkages among education and workforce programs and on how states use longitudinal data. In addition, GAO interviewed a nongeneralizable sample of five grantees, which were selected based on the progress they have made in matching data and on the funding they have received from the SLDS and WDQI programs. GAO also reviewed relevant federal laws and regulations.

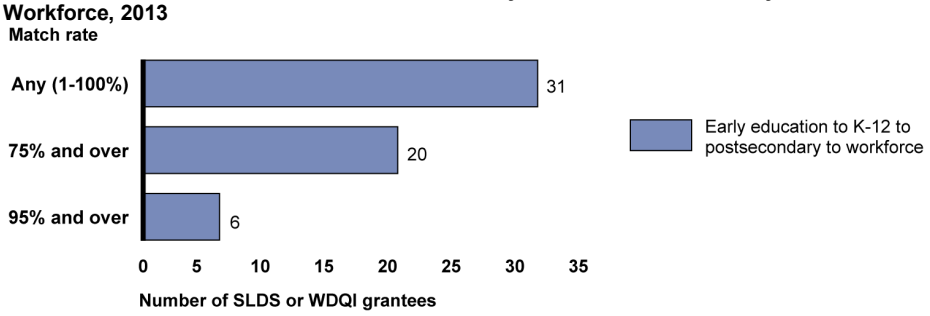
GAO is not making recommendations in this report. GAO received technical comments on a draft of this report from the Department of Education and the Department of Labor, and incorporated them as appropriate.

View GAO-15-27. For more information, contact Jacqueline M.Nowicki at (617) 788-0580 or nowickij@gao.gov.

What GAO Found

Over half of 48 grantee states that received a Statewide Longitudinal Data Systems (SLDS) or Workforce Data Quality Initiative (WDQI) grant have the ability to match data on individuals from early education into the workforce, based on GAO's analysis of 2013 Data Quality Campaign (DQC) survey data. The DQC is a nonprofit organization that supports the effective use of data to improve student achievement. In its survey, DQC collected self-reported information from states on their ability to match, or connect the same individual record, between the (1) K-12 and early education, postsecondary, and workforce sectors and between the (2) postsecondary and workforce sectors. However, as the match rate—that is, the percent of unique individual records reliably connected between databases—increases, the number of grantees able to match data decreases. GAO found that more grantees reported being able to match data among the education sectors than between the education and workforce sectors. Further, most grantees reported that they are not able to match data comprehensively. For example, only 6 of 31 grantees reported that they match K-12 data to all seven possible workforce programs covered by the DQC survey, which include adult basic and secondary education as well as unemployment insurance wage records. State officials cited several challenges to matching data, including state restrictions on the use of a Social Security number. Specifically, officials in three of five grantee states GAO spoke with said state law or agency policy prohibit collecting a Social Security number in K-12 data, which can make it more difficult to directly match individuals' K-12 and workforce records.

Number of SLDS or WDQI Grantees with the Ability to Match Data from Early Education to Workforce, 2013



Source: GAO analysis of 2013 Data Quality Campaign survey data. | GAO-15-27

According to GAO analysis of the DQC survey data, grantees use some longitudinal data to inform policy decisions and to shape research agendas. All 48 grantees reported analyzing aggregate-level data to help guide school-, district-, and state-level improvement efforts. For example, 27 grantees said they analyze data on college and career readiness to help schools determine whether students are on track for success in college or in the workforce. Grantees also reported using longitudinal data to analyze outcomes for individual students. For example, 29 grantees reported that they produce early warning reports that identify students who are most likely to be at risk of academic failure or dropping out of school. Data from the DQC survey also show that 39 grantees reported developing a research agenda in conjunction with their longitudinal data systems.

Contents

Figures

Abbreviations

America COMPETES Act	America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act
DQC	Data Quality Campaign
DOL	U.S. Department of Labor
Education	U.S. Department of Education
SSN	Social Security number
SLDS	Statewide Longitudinal Data Systems
TANF	Temporary Assistance for Needy Families
WDQI	Workforce Data Quality Initiative
WIA	Workforce Investment Act of 1998

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

November 19, 2014

The Honorable Tom Harkin
Chairman
Committee on Health, Education,
Labor, and Pensions
United States Senate

Dear Mr. Chairman:

Over the past decade, states have taken steps to create longitudinal data systems that will allow them to track individuals' participation in education and workforce programs over time. Having individual-level longitudinal data, such as information on individual characteristics and educational attainment, may help answer long-standing questions about the effectiveness of education and workforce programs in improving outcomes for students as they move through school and into the workforce, as well as for workers. These data could, for example, be used to evaluate teacher performance, identify school graduation rates, or determine the supply of skilled workers in relation to the demand for workers.

From fiscal years 2006 through 2013, the federal government provided over $640 million through two grant programs to support states' efforts to develop or enhance these longitudinal data systems. To date, the Department of Education's (Education) Statewide Longitudinal Data Systems (SLDS) program has awarded approximately $613 million to 47 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. The Department of Labor's (DOL) Workforce Data Quality Initiative (WDQI) grant program has awarded about $36 million to 33 states.[1]

You asked us to review the status of grantees' longitudinal data systems. This report, in turn, examines: (1) the extent to which SLDS and WDQI grantees match individual student and worker records and share data between the education and workforce sectors[2] and (2) how grantees are

[1] Some states received more than one grant.

[2] Matching is defined as reliably connecting the same individual record in two or more databases. Sharing is defined as exchanging data between two databases, in either direction. See our glossary of terms.

using longitudinal data to help improve education and workforce outcomes.

To answer these questions, we analyzed state-level data from a 2013 survey conducted by the Data Quality Campaign (DQC).[3] This survey collected self-reported information from 49 states and the District of Columbia on data linkages between the education and workforce sectors—early education,[4] K-12, postsecondary education, and workforce programs—and on specific programs within these sectors. In addition, the survey asked questions about how states analyze and use longitudinal data.[5] We reviewed the survey instrument, interviewed officials responsible for administering the survey, and tested the data for obvious inaccuracies and determined that these data are sufficiently reliable for the purposes of this report. Our data analysis was limited to the 48 states that (1) received a SLDS grant, a WDQI grant, or both and (2) responded to DQC's survey.[6] These 48 states were awarded grants from fiscal year 2006 through fiscal year 2013, the most recent year for which we had available data on grants awarded. In addition to our analysis of DQC survey data, we interviewed Education and DOL officials. We also conducted interviews with five grantees—Ohio, Pennsylvania, South Dakota, Virginia, and Washington—to obtain more information about their longitudinal data systems and, in particular, any challenges they faced in developing and implementing them. During these interviews, we also discussed how grantees have used longitudinal data to inform decision-making in education and workforce programs. We selected these grantees because they differed in terms of the progress they have made in establishing data linkages and on the federal funding they have

[3] DQC is a nonprofit organization that works with state officials and others to support the effective use of data to improve student achievement.

[4] For the purposes of this report, we use the term "early education" even though some of the programs discussed may go beyond education services. Various terms such as early childhood, early learning, or early care and education are also used to describe this sector.

[5] See appendix II for the specific questions we analyzed.

[6] The 48 states include the District of Columbia. For the purposes of this report, we will refer to the District of Columbia as a state. California chose not to participate in DQC's 2013 survey. We excluded Alabama and New Mexico from our analysis because neither state received a SLDS or WDQI grant. While Puerto Rico and the U.S. Virgin Islands received SLDS grants, we excluded them from our analysis because DQC did not survey these territories.

received from the SLDS and WDQI programs. Our findings from these interviews are not generalizable to all grantees, but they provide information on a range of challenges faced by grantees in developing and implementing longitudinal data systems, as well as some examples of how grantees are using longitudinal data. We also reviewed relevant federal laws and regulations. While we spoke to some grantees and reviewed grant requirements generally, we did not examine whether states have met the conditions set forth in their grant agreements. See appendix I for more information about our objectives, scope, and methodology.

We conducted this performance audit from October 2013 through November 2014 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Since the early 2000s, states have been building longitudinal data systems to better address data collection and reporting requirements in federal laws—such as the No Child Left Behind Act of 2001 and the America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act (America COMPETES Act)—and to inform stakeholders about student achievement and school performance.[7] Federal, state, and private entities have provided funding for these systems. For example, in addition to the SLDS and WDQI programs, other recent federal grant programs including Race to the Top and the Race to the Top-Early Learning Challenge may support states' efforts.[8]

[7] The No Child Left Behind Act of 2001 included provisions that required states to collect data and report on various indicators of student achievement, accountability, and teacher quality. The America COMPETES Act included requirements for longitudinal data systems of states receiving grants under that law, including the particular data elements the system should capture. 20 U.S.C. § 9871.

[8] States can use Race to the Top or Race to the Top-Early Learning Challenge program funds to implement a range of education reforms, including the development and implementation of their longitudinal data systems. We did not examine the extent to which states actually used funds from these programs to develop their longitudinal data systems.

The purpose of the SLDS grant program—administered by Education's Institute for Education Sciences, National Center for Education Statistics—is generally to enable state educational agencies to design, develop, implement, and expand statewide longitudinal data systems to manage, analyze, disaggregate, and use individual student data. From fiscal years 2006 to 2013, Education has awarded approximately $613 million in SLDS grants (see table 1).[9] For each grant competition, Education establishes the award period and range of grant amounts to be awarded; SLDS grants have ranged anywhere from 3 to 5 years, with a maximum award amount of $20 million per grantee. See appendix III for a list of states that received SLDS grants and the amount of their awards.

Table 1: Statewide Longitudinal Data Systems Grants Awarded, by Fiscal Year

Fiscal Year	Number of Grants Awarded	Total Funds Awarded (in millions)
2006	14	$52.8
2007	13	$62.2
2008	—	—
2009	27	$150.4
2009 ARRA[a]	20	$250.0
2010	—	—
2011	—	—
2012	24	$97.3
2013	—	—
Total	**98**	**$612.7**

Source: GAO analysis of Department of Education data. | GAO-15-27

Notes: Number of grants awarded includes grants made to Puerto Rico and the U.S. Virgin Islands in addition to states. Some states were awarded more than one SLDS grant from fiscal years 2006 to 2013. Education officials told us they chose not to award new grants in fiscal years 2008, 2010, 2011, and 2013 in order to provide larger awards in later years to support significant work on longitudinal data systems.

[a]Grants for the 2009 ARRA competition were awarded using funds from the American Recovery and Reinvestment Act of 2009, known as ARRA or the Recovery Act.

Though the SLDS grant requirements have varied over time, states generally could use SLDS funds to build K-12 longitudinal data systems

[9] Throughout this report, we refer to the various rounds of SLDS funding by fiscal year to be consistent with how Education refers to them.

or to expand these systems to include data from other sectors, such as early education, postsecondary education, or workforce (see table 2). The long-term goal of the program is for states to create comprehensive "P20-W"—early learning through workforce—longitudinal data systems that, among other things, will allow for states, districts, schools, educators, and other stakeholders to make informed decisions and conduct research to improve student academic achievement and close achievement gaps.

Table 2: Data Linkages under Statewide Longitudinal Data Systems Grants, by Fiscal Year

Fiscal Year	Required to Link K-12 Data to Other Sectors?	Grant Conditions for Linking K-12 Data to Other Sectors
2006	No	Focus of the grant was to build K-12 longitudinal data systems[b]
2007	No	States were encouraged to link K-12 data with postsecondary data
2009	Yes	States could propose to expand their existing K-12 data systems to include pre-k, postsecondary, or workforce data; if they did so, they were required to implement the proposed expansion(s)
2009 ARRA[a]	Yes	States were required to build systems that include data from preschool through postsecondary education and into the workforce
2012	Yes	States could apply under one of three grant priorities: (1) design, develop, and implement statewide, longitudinal K-12 data systems; (2) develop and link early childhood data to the state's K-12 data system; or (3) develop and link postsecondary and/or workforce data to the state's K-12 data system; states that applied under priorities (2) or (3) were required to implement those linkages

Source: GAO analysis of Department of Education Requests for Applications. | GAO-15-27

[a]Grants for the fiscal year 2009 ARRA competition were awarded using funds from the American Recovery and Reinvestment Act of 2009, known as ARRA or the Recovery Act.

[b]Education officials also mentioned exchanging data between secondary and postsecondary institutions as a potential example of the required capacity to exchange data across institutions within the state.

Under the WDQI grant program—administered by DOL's Employment and Training Administration—states are expected to fully develop their workforce longitudinal data systems and then be able to match these data with available education data to analyze education and workforce

outcomes.[10] This program complements Education's SLDS program. DOL has chosen to award WDQI grants to states that have received an SLDS grant or have a longitudinal data system in place. Among other requirements, all grantees are required to develop or improve workforce longitudinal data systems and enable workforce data to be matched with education data to ultimately follow individuals through school and into the workforce. DOL has provided funding for approximately $36 million in WDQI grants to 33 states since fiscal year 2010 (see table 3). The award period for each grant is 3 years. See appendix III for a list of states that received WDQI grants and the amount of their awards.

Table 3: Workforce Data Quality Initiative Grants Awarded, by Fiscal Year

Fiscal Year[a]	Number of Grants Awarded	Total Funds Awarded (in millions)
2010	13	$12.2
2011	12	$11.9
2012	6	$6.4
2013	6	$5.9
Total	37	$36.4

Source: GAO analysis of Department of Labor data. | GAO-15-27

Notes: Some states received more than one grant.

[a]Fiscal year refers to the timeframe of the applicable appropriation funding the grant competition.

While all grantees are expected to link education and workforce data, the expectations for doing so vary depending on the state's progress in developing its workforce longitudinal data system. More specifically, states without a workforce longitudinal data system are expected to enable their workforce systems to be linked to existing education data systems, states with partial workforce longitudinal data systems are expected to enable linkages to existing education longitudinal data systems, and states with comprehensive workforce longitudinal data systems are expected to improve linkages with education systems.

[10] The program's purposes, which have generally remained the same over time, include: (1) to develop or improve state workforce longitudinal data systems, (2) to enable workforce data to be matched with education data, (3) to improve the quality and breadth of data in workforce longitudinal data systems, (4) to use longitudinal data to evaluate the performance of education and job training programs, and (5) to provide user-friendly information to consumers to help them select education and training programs.

Over Half of Grantees Could Track Some Individuals from Early Education into the Workforce, but Data Are Generally Limited

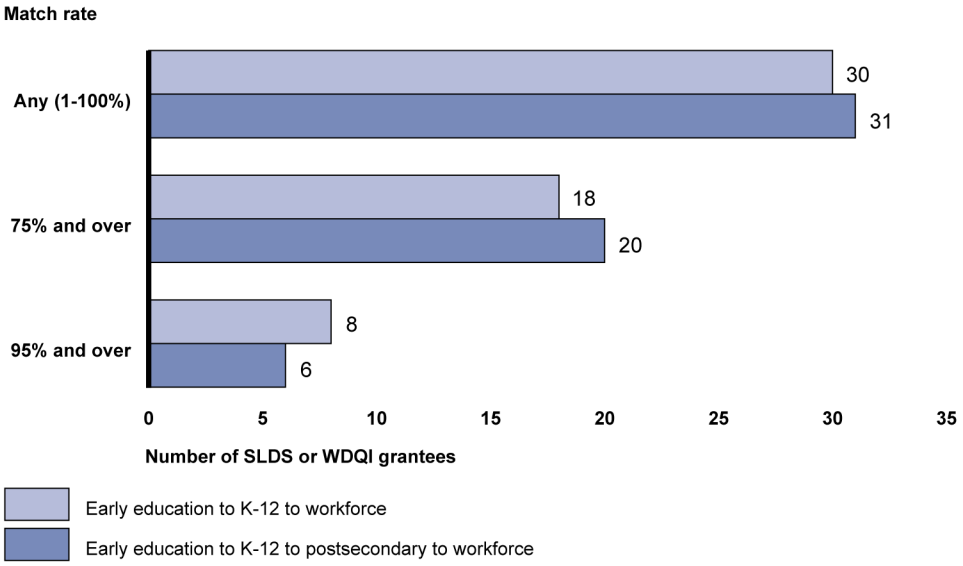
Over Half of Grantees Have the Ability to Match Some Individual Records from Early Education to Workforce

After analyzing data from DQC's 2013 survey,[11] we determined that over half of grantees have the ability to match data—reliably connect the same record in two or more databases—for some individuals from early education and into the workforce.[12] As shown in figure 1, individuals can take different paths to move from early education into the workforce: (1) via K-12 or (2) via K-12 and postsecondary. Regardless, as the match rate—that is, the percent of unique student records reliably connected between databases—increases, the number of grantees able to match data between sectors decreases. For example, 31 of 48 grantees have the ability to track individuals between all sectors from early education to workforce to at least some degree, but only 6 grantees could do so at the highest match rate.

[11] As noted previously, these data are self-reported by states to DQC.

[12] For the purposes of this report, a grantee is one of the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 Data Quality Campaign survey. DQC's data do not allow us to determine whether records for the same individual are actually being matched between all sectors from early education to workforce. The data only allow us to determine whether a grantee has the *ability* to match data between multiple sectors, which we determined by analyzing grantees' responses to individual survey questions about their ability to match data between K-12 and early education, K-12 and postsecondary, K-12 and workforce, and workforce and postsecondary. If a grantee reported, for example, being able to match data on one program between K-12 and early education and on one program between K-12 and workforce, we concluded that the grantee has the ability to match data from early education to workforce.

Figure 1: Number of SLDS or WDQI Grantees with the Ability to Match Data from Early Education to Workforce, 2013



Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed data on the 48 states that received a State Longitudinal Data Systems (SLDS) grant, a Workforce Data Quality Initiative (WDQI) grant, or both and responded to the 2013 DQC survey. Matching is defined as reliably connecting the same individual record in two or more databases. The match rate is the percent of unique individual records that are matched. We considered a grantee as matching data between sectors if a grantee matched data from at least one program between sectors.

Our analysis of the DQC survey data also shows that more grantees match data among the education sectors than between the education and workforce sectors, though—as was the case with matching data from early education to workforce—the number of grantees that match data decreases as the match rate increases (see table 4).[13] For example, 43 grantees reported matching data between the K-12 and early education sectors, and 31 grantees reported matching data between the K-12 and workforce sectors at least to some degree; however, the number of grantees that reported matching data between these same sectors drops to 37 and 9, respectively, at a match rate of 95 percent or more.

[13] We considered a grantee as matching data across sectors if a grantee matched data from at least one program across sectors.

Table 4: Number of SLDS or WDQI Grantees Matching Data between Sectors, 2013

Match Rate	K-12 and Early Education	K-12 and Postsecondary	K-12 and Workforce	Postsecondary and Workforce
Any (1-100%)	43	43	31	36
75% and Over	42	35	19	25
95% and Over	37	15	9	14

Source: GAO Analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed data on the 48 states that received a State Longitudinal Data Systems (SLDS) grant, a Workforce Data Quality Initiative (WDQI) grant, or both and responded to the 2013 DQC survey. Matching is defined as reliably connecting the same individual record in two or more databases. The match rate is the percent of unique individual records that are matched. We considered a grantee as matching data between sectors if a grantee matched data from at least one program between sectors.

Not all grantees are matching data between all sectors, which may partially be the result of receiving grants with different grant requirements. For example, all 20 grantees that received a fiscal year 2009 SLDS ARRA grant were required to have longitudinal data systems that include individual student-level data from preschool through postsecondary education and into the workforce (see table 2). However, fiscal year 2012 grantees could choose from among three different grant priorities, so some grantees may be focused on building a K-12 longitudinal data system while others may be using their grant funds to link existing K-12 data to other sectors. In addition, grantees may have been in different stages of developing their longitudinal data systems prior to receiving a grant, which may help explain why some grantees are able to match data between more sectors than others.

Most Grantees Are Not Able to Match Data Comprehensively

Programs Included in the Data Quality Campaign Survey, 2013

Early education: early intervention, Head Start/Early Head Start, special education, state prekindergarten, subsidized child care

K-12: elementary and secondary education

Postsecondary institutions: less than 2-year public, less than 2-year private not-for-profit, less than 2-year private for-profit, 2-year public, 2-year private not-for-profit, 2-year private for-profit, 4-year and above public, 4-year and above private not-for-profit, 4-year and above private for-profit

Workforce: unemployment insurance wage records, unemployment benefits claim data, Workforce Investment Act of 1998 (WIA) adult or dislocated worker program, WIA youth program, adult basic and secondary education, Wagner-Peyser Act employment services, Temporary Assistance for Needy Families (TANF)

Source: 2013 Data Quality Campaign Survey (DQC) Instrument. | GAO-15-27

Of those grantees that match data, we found that few generally do so for all of the possible programs between particular sectors (see sidebar),[14] based on our analysis of DQC survey data (see table 5). For example, only 6 of 31 grantees reported that they were able to match data on all seven programs between the K-12 and workforce sectors, which include unemployment insurance wage records, unemployment benefit claims data, Workforce Investment Act of 1998 (WIA) adult or dislocated worker program, WIA youth program, adult basic and secondary education, Wagner-Peyser Act employment services, and Temporary Assistance for Needy Families (TANF).

Table 5: Number of Programs Matched by Grantees, Any Match Rate, 2013

Number of Programs Matched	K-12 and Early Education	K-12 and Postsecondary	K-12 and Workforce	Postsecondary and Workforce
One	0	0	10	10
Some[a]	35	33	15	20
All	8	10	6	6
Total Number of Grantees that Match at Any Match Rate (1-100%)	**43**	**43**	**31**	**36**

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

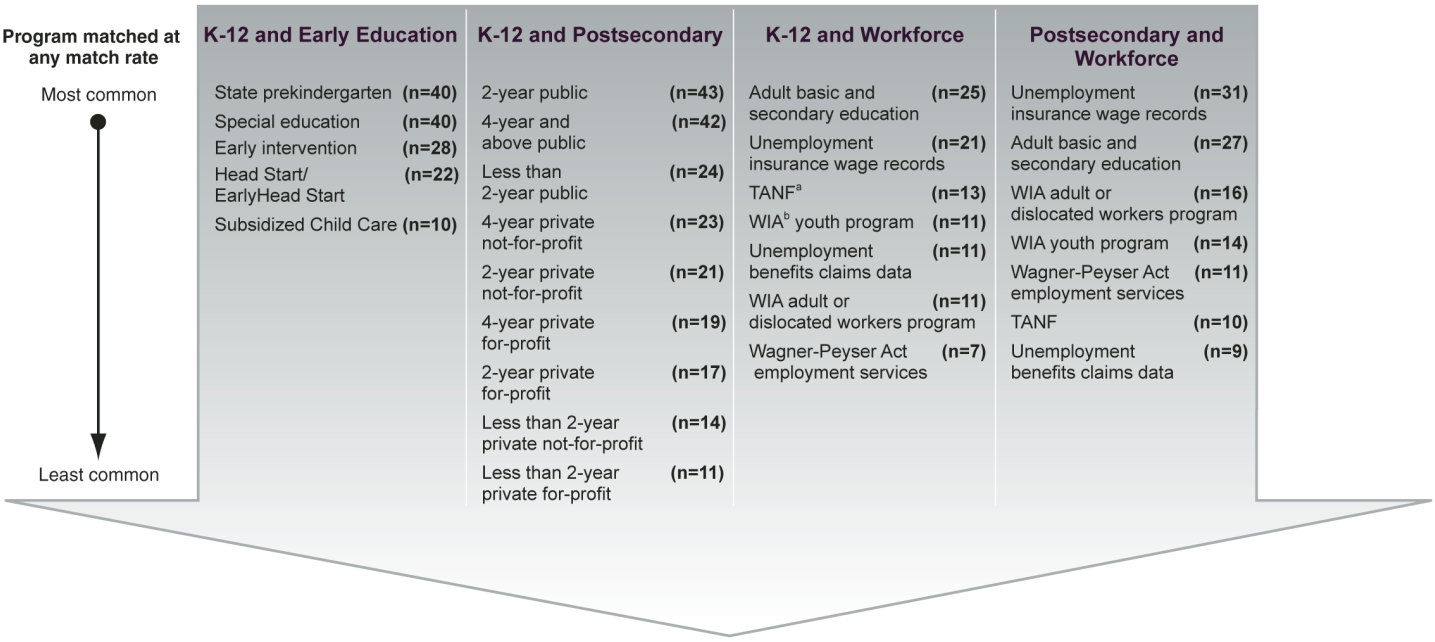
Notes: We analyzed data on those grantees that match data between particular sectors at any match rate (see table 4). We define a grantee as any of the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 DQC survey. DQC asked states to identify the programs for which they match data between sectors. For a list of programs, see the definitions of each sector included in our glossary of terms.

[a]We define some as more than one but less than all possible programs that could be matched between sectors.

We also analyzed DQC's data to determine which programs are most commonly matched by grantees between particular sectors (see fig. 2). See appendix IV for a list of the specific programs matched by each grantee.

[14] For the purposes of this report, we refer to types of postsecondary institutions as programs.

Figure 2: Most to Least Commonly Matched Programs between Sectors, Any Match Rate, 2013



Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed the types of programs matched by those grantees that indicated that they match data between particular sectors at any match rate (see table 4). We define a grantee as any of the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 DQC survey. The n values represent the number of states that reported matching data on a particular program; some programs have the same n value. The figure is not to scale.

[a]TANF is the Temporary Assistance for Needy Families program.

[b]WIA is the Workforce Investment Act of 1998.

Most Grantees That Match Data Also Share Them, but Few Share All Types of Data

Most grantees that match data also share data between sectors; that is, they exchange at least one type of data (e.g., demographic, enrollment, program participation, etc.) between two databases in at least one direction, based on our analysis of DQC data. However, in general, few grantees share all possible types of data (see sidebar). For example, only 3 of 36 states that match data between the postsecondary and workforce sectors reported sharing all 10 types of data asked about by DQC, which include information on postsecondary degree completion, earnings and wages, and industry of employment, among others (see table 6).

Types of Data Included in the Data Quality Campaign Survey, 2013

Early education: demographic, child-level developmental data, family characteristics, program participation

K-12: child-level developmental data, college readiness assessment scores, demographic, enrollment, family characteristics, financial aid, transcript data, program participation

Postsecondary: college placement assessment scores, demographic, enrollment, financial aid, postsecondary credits earned, postsecondary degree completion, postsecondary enrollment, postsecondary enrollment intensity, postsecondary outcomes, postsecondary progress, postsecondary remediation status, postsecondary type of degree, transcript data

Workforce: demographic, earnings and wages, employment status, enrollment, industry of employment, occupation

Source: 2013 Data Quality Campaign Survey (DQC) Instrument. | GAO-15-27

Table 6: Number of Grantees That Share Data between Sectors, Any Match Rate, 2013

Data Elements Shared	K-12 and Early Education	K-12 and Postsecondary	K-12 and Workforce	Postsecondary and Workforce
None	1	0	4	3
Some[a]	27	37	25	30
All	15	6	2	3
Total Number of Grantees that Match at Any Match Rate (1-100%)	**43**	**43**	**31**	**36**

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed the types of data shared by grantees that match data between particular sectors at any match rate (see table 4). We define a grantee as any of the 48 states that received a SLDS grant, a WDQI grant, or both and responded to DQC's 2013 survey. DQC asked states about the types of data shared between sectors. For a list of data elements DQC asked about, see the definitions of sharing included in our glossary of terms.

[a]We define some as at least one but less than all possible data elements that could be matched across sectors.

Absence of a Social Security Number in Education Data and Data Governance Were among Challenges to Matching Data

Officials in all five grantee states we spoke with said matching K-12 education and workforce data is challenging without using a Social Security number (SSN) that uniquely identifies an individual and, as a result, some states may have greater difficulty tracking particular groups of students over time. SLDS officials in three states—Ohio, Pennsylvania, and Virginia—said collecting a SSN in K-12 education data is prohibited either by state law or agency policy; in the other two states—South Dakota and Washington—officials said collecting a SSN is optional and whether to do so is determined at the district level.[15] While establishing a unique statewide student identifier is a technical requirement of the SLDS grant program, states can choose the format of the identifier used. Education suggested, in a November 2010 SLDS Technical Brief, that states use a unique identifier distinct from a student's SSN for privacy reasons; however, Education also stated that states should maintain a

[15] For the purposes of this report, we refer to state officials we spoke with about the SLDS grant or the WDQI grant as SLDS officials or WDQI officials, as appropriate.

student's SSN as a data element in order to link data between systems.[16] According to a 2010 report from the Social Security Administration's Office of the Inspector General, 28 states collect a SSN in K-12 education data.[17]

Unlike the SLDS program, in its evaluation criteria for WDQI grants, DOL specifies that states use SSNs as a personal identifier, as they are already in use throughout the workforce system. To match education and workforce data absent a SSN, state officials said they are developing algorithms to match individual records using other identifiers, which could include an individual's first name, last name, and date of birth. However, a person's last name can change, which Pennsylvania SLDS officials said can make it difficult to reliably track individuals over time. Further, Ohio WDQI officials explained that the absence of a SSN makes it particularly difficult to track students who drop out of high school or to track high school graduates who do not move on to the workforce. Similarly, Ohio SLDS officials said tracking students that do not go on to postsecondary education is a challenge because there is no readily available identifier to determine any workforce participation by those individuals.

In four of five grantee states we spoke with, officials also cited data governance as a challenge. Data governance is the exercise of decision-making and authority for data-related matters using agreed-upon rules that describe who can take what actions with what information and when, under what circumstances, and using what methods.[18] SLDS grantees are generally required to develop a governance structure involving both state and local stakeholders that includes a common understanding of data ownership, data management, as well as data confidentiality and access. All WDQI grantees are expected to establish partnerships with relevant workforce agencies and with state education agencies for the

[16] U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, *Data Stewardship: Managing Personally Identifiable Information in Electronic Student Education Records*, NCES 2011-602 (Washington, D.C.: November 2010).

[17] Social Security Administration, Office of the Inspector General, *Kindergarten Through 12th Grade Schools' Collection and Use of Social Security Numbers*, A-08-10-11057 (Washington, D.C.: July 2010).

[18] This definition is based on information from the Data Governance Institute (DGI). According to its website, DGI provides vendor-neutral data governance practices and guidance. See www.datagovernance.com for more information.

purposes of data sharing. Pennsylvania and Ohio officials said it has not been easy to get the various workforce agencies that maintain data on individual workforce programs to share their data as the agencies often operate independently from one another. As a result, Pennsylvania officials said agencies are territorial about their data, making it difficult to build consensus around developing a longitudinal data system. In Ohio, officials said that each agency has to be approached separately to obtain commitment to share data in a longitudinal system. Similarly, officials in Virginia said collecting data on early education programs has been a challenge as the data are scattered across different agencies. An official from the Early Childhood Data Collaborative explained that it can be easier to facilitate data matching between early education programs under the purview of one agency, such as state prekindergarten and special education, which are generally overseen by state educational agencies in addition to K-12 data.[19]

Based on our interviews with grantee states, state officials we spoke with said they are in different stages of developing a data governance structure. For example, Pennsylvania WDQI officials said they have not yet established a formal data governance structure. In contrast, Virginia officials have established a data governance structure; officials said they spent 18 months working through the different priorities, cultures, and agendas of the various agencies providing data to the longitudinal data system.

State officials in all five grantee states we spoke with also said they have had to manage public concerns about the purpose of data collection or about data privacy. For example, in Ohio, SLDS officials told us there is a lack of understanding about the value of building a longitudinal data system; officials have had to counter misperceptions about what data are being collected in the state's longitudinal data system, what the data will be used for, and why data need to be connected between the education and workforce sectors. South Dakota officials said they have had to respond to concerns from parents and other education stakeholders about the privacy of longitudinal data.

[19] The Early Childhood Data Collaborative supports state policymakers' development and use of coordinated state early care and education data systems.

Grantees have tried to provide information to the public about the purposes of the data system and steps taken to safeguard information. Forty-six grantees reported using outreach tools to communicate the availability of the data to non-educator stakeholders,[20] according to our analysis of the DQC survey data. These grantees reported using traditional outreach measures, which could include public service announcements, press conferences and news releases, and posting information about the data on the state education agency's website. For example, four of five grantee states we interviewed have web pages dedicated to their longitudinal data systems. These web pages can include overviews of the systems, answers to frequently asked questions, trainings on how to use or access the data, and examples of research studies that use the data. Further, 44 grantees reported on the DQC survey that they take advantage of in-person opportunities, which could include meetings, conferences, and presentations. Lastly, 35 grantees reported using electronic or social media to promote the data, which could include Facebook, Twitter, blogs, and webinars. In the context of discussing the challenge of managing public concerns about data collection or privacy, officials in three of the five grantee states we spoke with specifically said they have provided information about how they protect individual data. Pennsylvania SLDS officials said they took considerable time to convey to parents and taxpayers the steps they are taking to ensure data privacy. Similarly, Virginia officials from both grant programs said explaining all of the precautions the state is taking with respect to data privacy seems to help in reducing concerns. Ohio officials said the state's Department of Education has convened a new workgroup to see if there are better ways to address misperceptions about data collection and use.

Lastly, state officials cited the importance of federal funding to their efforts to build their longitudinal data systems and expressed concerns about sustaining their systems after their grants end. Officials we interviewed in all five grantee states said they would not be as far along in developing their longitudinal data systems without the federal funding provided through the SLDS and WDQI programs. For example, officials in Washington said they used their initial SLDS and WDQI grants to focus on building their K-12 data system and workforce systems, respectively.

[20] The DQC defined non-educator stakeholders to be stakeholders who are not teachers, principals, or superintendents.

They said the second SLDS grant they received was instrumental in building a P-20W system to connect data between all sectors. Ohio officials said the SLDS funds have provided, among other uses, critical funding for further development of the longitudinal data system, technological updates, and access to technical assistance. However, officials in all five grantee states also expressed concerns about sustaining the systems moving forward. For example, officials in Virginia said they have created a legislative committee to focus on sustainability efforts and will need to request additional funding to keep the system sustainable. Officials in Pennsylvania said they are trying to leverage the existing technical infrastructure and use other available resources, but it is difficult to find funding for their workforce data efforts.

Longitudinal Data Analysis Has Informed State and Local Policy Making and Helped States Shape Their Research Agendas

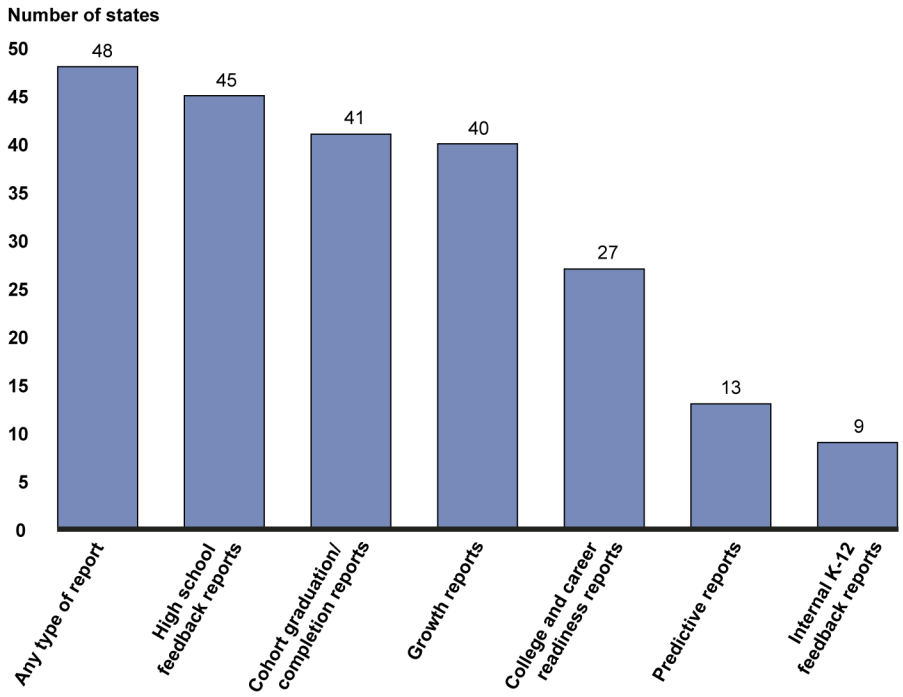
Grantees Reported Using Some Longitudinal Data to Inform Decision and Policy Making

According to our analysis of the DQC survey data and our interviews with selected states, SLDS and WDQI grantees use longitudinal data to examine education outcomes and to inform policy decisions. All 48 grantees responded that their state educational agency uses the data to analyze aggregate education outcomes (see fig. 3).[21] For example, the three most common types of analyses are related to high school feedback, cohort graduation or completion, and growth (i.e., changes in the achievement of the same students over time). These aggregate data are used to analyze a particular cohort of students and develop information on students' outcomes over time. They also help guide school-, district-, and state-level improvement efforts. For example,

[21] The DQC defined aggregate level-data to be group statistics, such as numbers, percentages, and averages, based on individual student data.

officials from three of the five grantee states we interviewed told us they have used the data to assess kindergarten readiness for children who attended state early education programs. Also, 27 grantees responded to the DQC survey that they use the data to analyze college and career readiness. More specifically, to better understand the courses and achievement levels that high school graduates need to be successful in college, Virginia followed students who graduated from high school from 2006 to 2008 and analyzed enrollment and academic achievement patterns for different groups of students. According to agency officials in Virginia, this analysis resulted in changes to the course requirements for graduation.

Figure 3: Grantees' Use of Longitudinal Data Analysis to Examine Education Outcomes, 2013



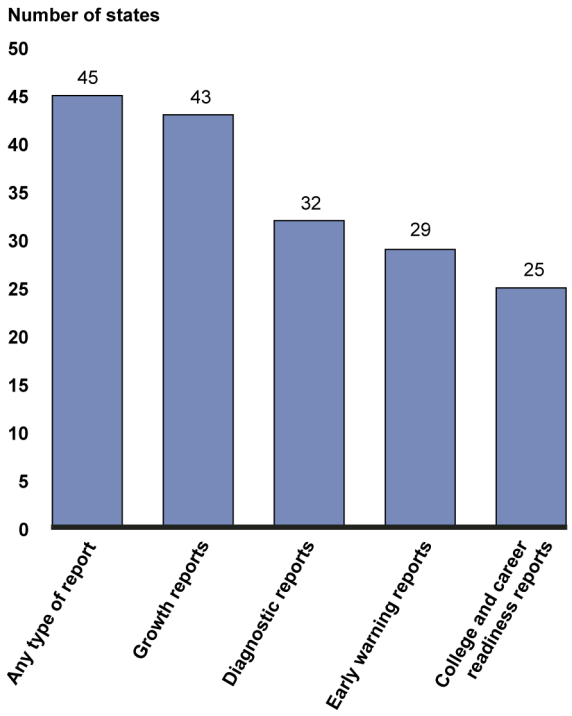
Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed data on the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 DQC survey. States may create more than one type of report. For the survey question and a complete list of response options, see appendix II. For definitions of the report types, see our glossary of terms.

In addition to examining education outcomes, states also use longitudinal data to assess how cohorts of students fare once they are in the workforce. Washington's Education Research and Data Center, a state center dedicated to analyzing education and workforce issues across the P-20W spectrum, has published several studies examining workforce outcomes for high school and college graduates. For example, one study compared earnings for workers with bachelor's degrees from Washington state colleges and universities to earnings of workers with only diplomas from public high schools.

In addition to analyzing aggregate student outcomes, grantees also indicated that they analyze individual-level student outcomes. Our analysis of DQC survey data shows that 45 of 48 grantees examine outcomes for individual students (see fig. 4). Student-level data provide teachers and parents with information they can use to improve student achievement. For example, 32 grantees reported that the data are used in diagnostic analysis, which help teachers identify individual students' strengths and academic needs. Also, 29 grantees responded to the DQC survey that they produce early warning reports, which identify students who are most likely to be at risk of academic failure or dropping out of school. For example, Virginia's early warning report shows demographic and enrollment information about an individual student; flags for warning indicators such as attendance, GPA, and suspensions; and a record of interventions the school has taken to help the student (see fig. 5). Further, officials in three of the grantee states we interviewed told us that educators have access to student-level analyses. In Pennsylvania, teachers can use an educator dashboard, which includes longitudinal data, to determine the educational needs of their students and adjust their teaching plans.

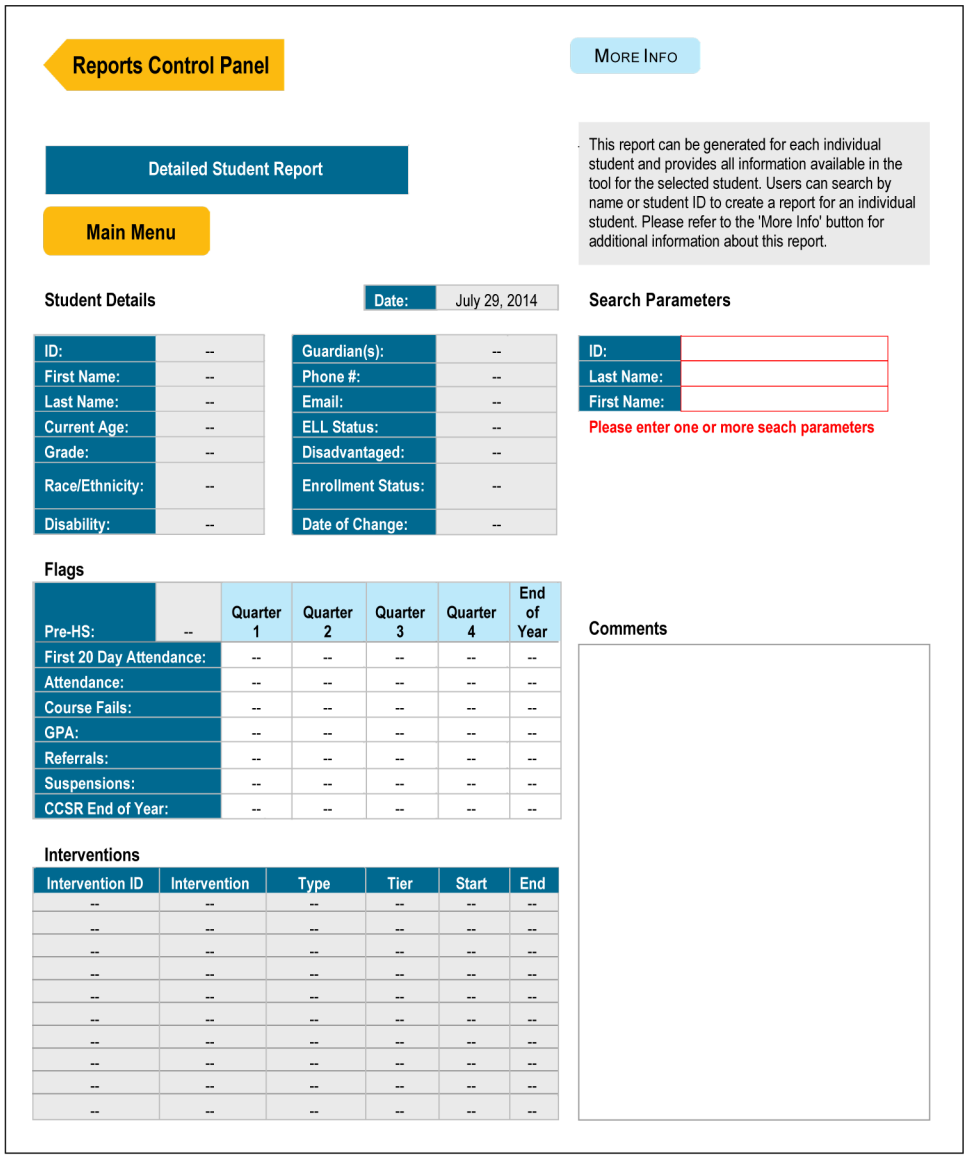
Figure 4: Grantees' Use of Longitudinal Data Analysis to Examine Outcomes for Individual Students, 2013



Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: We analyzed data on the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 DQC survey. States may create more than one type of report. For the survey question and a complete list of response options, see appendix II. For definitions of the report types, see our glossary of terms.

Figure 5: Example of an Early Warning Report from Virginia



Source: Virginia Department of Education. | GAO-15-27

Forty-one of 48 grantees reported to the DQC that they use longitudinal data to inform policy and continuous improvement efforts. Specifically, grantees reported that they use the data to inform school turnaround efforts (34 grantees), evaluate intervention strategies or programs (14 grantees), or identify and reward schools that demonstrate high growth

(27 grantees), among other things. Officials in three of five grantee states we spoke with provided more specific examples of how they use or plan to use longitudinal data to inform their efforts. Ohio officials told us they used longitudinal data to study students in remediation to help develop a remediation policy. They also said they have been working on a workforce success measures dashboard to compare outcomes across state programs. For example, the dashboard will allow policy makers to assess how successful the state's adult basic education program is compared to the state's vocational education program. Pennsylvania officials told us they will develop a similar dashboard. Washington state officials told us that longitudinal data helped address a concern in the state legislature about whether math and science teachers were leaving to work in the private sector. Researchers identified common teacher and school district characteristics associated with teachers who left for employment in other fields and found that math and science teachers did not leave the field at a higher rate than other teachers. Officials told us that this analysis prompted the state legislature to focus its attention on improving the recruitment of math and science teachers rather than improving retention.

While many grantees reported on the DQC survey that they use longitudinal data to analyze outcomes for students and workers and to make policy decisions, officials from all five grantee states we interviewed told us that these analyses are limited because they are still developing their longitudinal data systems. In addition, only three of these states—Ohio, Virginia, and Washington—are conducting education to workforce analyses. Officials in Pennsylvania and South Dakota said they plan to do this type of analysis, but only after they finish putting all the education and workforce data into their systems and matching these data.

Most Grantees Have Developed a Research Agenda in Conjunction with Their Longitudinal Data Systems

Data from the 2013 DQC survey show that 39 SLDS or WDQI grantees have developed research agendas articulating and prioritizing research or policy questions that can be answered with longitudinal data. These research agendas were developed in partnership with higher education institutions, independent researchers, or others. Of the five grantee states we interviewed, only Virginia and Ohio have fully developed their research agendas. Pennsylvania, South Dakota and Washington officials told us they are in the process of doing so. State officials shared two approaches for creating these agendas. Under the first approach, stakeholders from various state agencies comprise a committee that identifies research questions. Virginia took this approach and drafted a list of "burning questions" to answer using longitudinal data. Officials in

Virginia explained that they purposefully kept the agenda broad so that the questions will remain relevant over the long term. Washington's Education Research and Data Center has similarly developed a list of critical questions it would like to answer using longitudinal data. Under the second approach, state agencies use information requests and stakeholder feedback on sample reports to shape the research agenda. For example, officials from the South Dakota Department of Education told us they have solicited feedback after training districts on the data and reviewed requests from the governor's office and state legislators. They also told us that they are following the number of hits for individual reports on the state's Department of Education's electronic portal.

Forty-three of 48 grantees reported that they have a process by which researchers who are not employees of the state can propose their own studies for approval, according to the 2013 DQC survey data. Four of the grantee states we interviewed have established a formal request process for researchers who would like to access longitudinal data and the fifth state is reviewing its protocols and expects to develop a formal application process. Officials in two grantee states told us that the request process is intended to streamline access to the data and make it easier for researchers to seek approval for data requests. In addition, officials in Ohio told us that when researchers apply for access to Ohio's data, they must include information in their application about how the study will meet the state's research priorities.

Concluding Observations

Since fiscal year 2006, the federal government has made a significant investment—over $640 million in SLDS and WDQI grant funds—to help states build P20-W longitudinal data systems that track individuals from early education and into the workforce. The different grant requirements for linking data between sectors may have contributed to states being in different stages of developing their longitudinal data systems. That is, some grantees are just building their K-12 longitudinal data systems while others are matching data between education and workforce sectors. It remains to be seen whether all grantees will ultimately achieve the long-term goal of developing complete P20-W longitudinal data systems or how long that will take, particularly in light of unresolved concerns about limitations to matching data using a Social Security number and sustainability. Further, even among those grantees that can match data between sectors, most can only do so for a limited number of programs or data types. As grantees continue to refine their systems, maximizing the potential of these systems will rest, in part, with the ability to more fully

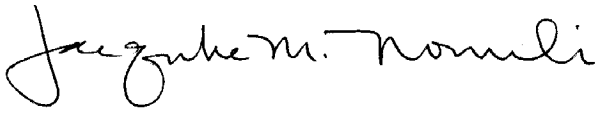
match information on specific programs and characteristics of individuals that could help in further analyzing education and workforce outcomes.

Agency Comments and Our Evaluation

We provided a draft of this report to Education and DOL for their review. Each provided technical comments, which we incorporated as appropriate.

As agreed with your office, unless you publicly announce its contents earlier, we plan no further distribution until 30 days after the date of this letter. At that time, we will send copies of this report to the appropriate congressional committees and the Secretaries of Education and Labor. In addition, the report is available at no charge on GAO's website at http://www.gao.gov. If you or your staff have any questions about this report, please contact me at (617) 788-0580 or nowickij@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix VII.

Sincerely yours,



Jacqueline M. Nowicki
Director, Education, Workforce,
and Income Security Issues

Appendix I: Objectives, Scope, and Methodology

The objectives of this report were to examine: (1) the extent to which Statewide Longitudinal Data Systems (SLDS) and Workforce Data Quality Initiative (WDQI) grantees match individual student and worker records and share data between the education and workforce sectors; and (2) how grantees are using longitudinal data to help improve education and workforce outcomes.

To answer our objectives, we analyzed state-level data from a 2013 survey conducted by the Data Quality Campaign (DQC), a nonprofit organization that works with state officials and others to support the effective use of data to improve student achievement. DQC's survey focused on 10 "State Actions" the DQC has developed to ensure effective data use (see table 7). DQC has conducted this annual survey since 2005.[1] The survey data include self-reported information on how data are matched and shared between the early education, K-12, postsecondary education, and workforce sectors, as well as information on specific programs within these sectors, how states analyze and use the data, and who has access to the data.

To conduct the survey, DQC used an online tool to collect information and invited the governor's office in all 50 states and the District of Columbia to participate. According to DQC, the governor's office is in the best position to bring stakeholders together to respond to the survey. As part of their survey response, states were asked to provide documents or website links as evidence of having specific policies or reports. After survey responses were received, DQC worked with each state to ensure the information reported was as accurate as possible.

[1] While DQC has conducted its annual survey since 2005, the survey questions have changed over time. At first, the survey questions were focused on determining if states had 10 "Essential Elements." DQC officials told us these 10 elements were 10 critical pieces states should have in a longitudinal data system. According to DQC officials, in 2009, the survey began asking states questions to determine whether they had implemented the 10 State Actions. DQC officials said the actions focus more on what states are doing with the data and whether the right people have secure access to the data.

Table 7: Data Quality Campaign's 10 State Actions to Ensure Effective Data Use

Action number	Action
1	Link state K-12 data systems with early learning, postsecondary, workforce, and other critical state agency data systems
2	Create stable, sustainable support for longitudinal data systems
3	Develop governance structures to guide data collection and use
4	Build state data repositories
5	Provide timely, role-based access to data while protecting privacy
6	Create progress reports with student-level data for educators, students, and parents
7	Create reports with longitudinal statistics to guide system-level change
8	Develop a purposeful research agenda
9	Implement policies and promote practices to build educators' capacity to use data
10	Promote strategies to raise awareness of available data

Source: Data Quality Campaign. | GAO-15-27.

We analyzed data from eight survey questions (see table 8 in appendix II) to determine the extent to which SLDS and WDQI grantees match individual records and share data among the education sectors and between the education and workforce sectors. For the purposes of our report, a grantee is one of the 48 states that received a SLDS grant, a WDQI grant, or both and responded to the 2013 DQC survey. We considered the District of Columbia to be a state. We excluded Alabama, New Mexico and California from our review because neither Alabama nor New Mexico received a SLDS or a WDQI grant and because California chose not to participate in DQC's 2013 survey. We excluded the U.S. Virgin Islands and Puerto Rico because, while these territories received SLDS grants, DQC did not include them in its survey. We analyzed data on SLDS and WDQI grantee states because the SLDS and WDQI grant programs provide federal funds for developing longitudinal data systems and are complementary.

We considered a grantee as matching data between sectors if a grantee matched data from at least one program between sectors (for a list of programs included in the DQC survey, see questions 1, 4, 7, and 10 in table 8 in appendix II). We considered a grantee as sharing data if a grantee matched data according to our definition and also reported exchanging at least one data element between sectors, in either direction

(for a list of data elements, see questions 2, 5, 8, and 11 in table 8 in appendix II). We also analyzed data from another twelve survey questions to identify how grantees are using longitudinal data to help improve education and workforce outcomes (see table 9 in appendix II).

We conducted a data reliability assessment by reviewing the survey instrument and related documentation, interviewing officials responsible for administering the survey, and testing the data for obvious inaccuracies. We determined that these data are sufficiently reliable for the purposes of this report.

In addition to our analysis of DQC survey data we conducted interviews with a nongeneralizable sample of five grantees as well as relevant federal agencies and nonprofit organizations. During our interviews with the five grantee states—Ohio, Pennsylvania, South Dakota, Virginia, and Washington—we asked grantees to identify challenges they faced in building and implementing longitudinal data systems and discussed how grantees have used longitudinal data to inform decision-making in education and workforce programs. We selected these grantees based on factors including the differing levels of progress they have made in establishing data linkages and the federal funding they have received from the SLDS and WDQI programs.

Within each state, we spoke with relevant K-12, workforce, postsecondary education, and early education officials. We also interviewed officials at Education, DOL, and the Department of Health and Human Services to obtain information about their roles in helping states build longitudinal data systems. In addition, we spoke with officials from nonprofit organizations to obtain their views on states' implementation of longitudinal data systems. These stakeholder organizations included the Early Childhood Education Collaborative, the State Higher Education Executive Officers Association, and the Workforce Data Quality Campaign.[2] Finally, we reviewed relevant federal laws, regulations,

[2] The Early Childhood Data Collaborative supports state policymakers' development and use of coordinated state early care and education data systems to improve the quality of early childcare and education programs and the workforce. The State Higher Education Executive Officers Association is a national association of state higher education leaders who serve statewide coordinating and governing boards and other state policy agencies for higher education. The Workforce Data Quality Campaign is a nonprofit, nonpartisan initiative that advocates for inclusive, aligned and market-relevant data systems used for advancing the nation's skilled workforce and helping U.S. industries compete in a changing economy.

requests for applications, and solicitations for grant applications to understand the requirements of these grants.

Appendix II: Data Quality Campaign (DQC) Survey Questions

As explained in appendix I, we analyzed data from DQC's 2013 survey to answer our research objectives. Table 8 and table 9 show the specific questions we analyzed from DQC's survey instrument.[1] For some questions, DQC allowed states to select "other" as a response; we excluded these "other" responses from our analysis.

Table 8: DQC Survey Questions We Analyzed to Answer Objective 1

Question number	Question wording	Response options
1	For each of the following early childhood programs, please indicate the range of the match rate when individual students' early childhood records are matched with their respective K–12 records at least annually. (Matching means to reliably connect the same student record based on one or more types of variables, including date of birth, first name, last name, middle name, student identifier, gender, etc.) 1. Subsidized child care 2. Early intervention 3. Special education 4. State prekindergarten 5. Head Start/Early Head Start	1. Do not match (0%) 2. 1-74% 3. 75-94% 4. 95-100% 5. Unknown match rate/have not analyzed
2	For the following types of student-level data, in which direction does data move between systems? 1. Demographic 2. Family characteristics 3. Program participation (e.g., bilingual/ESL, gifted and talented, migrant, special education, free- or reduced-price lunch) 4. Child-level developmental data	1. Early childhood to SEA 2. SEA to early childhood 3. This student-level information is NOT shared

[1] For more information about DQC's survey, see http://www.dataqualitycampaign.org/your-states-progress/about-data-for-action/.

Question number	Question wording	Response options
4	For each of the following types of postsecondary institutions, please indicate the range of the match rate when individual students' K–12 records are matched with their respective postsecondary records at least annually. Note that less than two-year institutions can include non-credit bearing institutions and technical colleges. 1. Less than two-year public 2. Less than two-year private not-for-profit 3. Less than two-year private for-profit 4. Two-year public 5. Two-year private not-for-profit 6. Two-year private for-profit 7. Four-year and above public 8. Four-year and above private not-for-profit 9. Four-year and above private for-profit	1. Do not match (0%) 2. 1-74% 3. 75-94% 4. 95-100% 5. Unknown match rate/have not analyzed
5	For the following types of student-level data, in which direction does data move between systems? Postsecondary systems refer to state higher education governing or coordinating bodies. Moving between systems means that data are moving between a state data system at the K–12 level (SEA) and one or more systems at the postsecondary level. Note that some options are not applicable (e.g., high school transcript data move from the SEA to postsecondary, not from postsecondary to the SEA). If you select an option that is not applicable in the online system, you will receive an error message. Select an applicable option in the online system to clear the error message. 1. Demographic 2. College readiness assessment scores (e.g., state, SAT, ACT, Advanced Placement) 3. College placement assessment scores (e.g., Accuplacer, Compass or other tests from postsecondary) 4. High school transcript data (e.g., course enrollment, grades, success) 5. Postsecondary enrollment 6. Postsecondary remediation status (e.g., remediation or developmental courses) 7. Postsecondary progress (e.g., success in first-year courses, retention, time and credit to degree, course completion) 8. Postsecondary credits earned 9. Postsecondary enrollment intensity (i.e., full time, part time) 10. Postsecondary outcomes (e.g., transfer, completion status, degree or certificate) 11. Financial aid	1. SEA to postsecondary 2. Postsecondary to SEA 3. This student-level information is NOT shared

Question number	Question wording	Response options
7	For each of the following employment programs or types of employment information, please indicate the range of the match rate when individual students' K–12 records are matched with their respective workforce records at least annually. 1. Unemployment insurance wage records 2. Unemployment benefits claim data 3. WIA adult or dislocated worker program 4. WIA youth program 5. Adult basic and secondary education 6. Wagner-Peyser Act employment services 7. TANF	1. Do not match (0%) 2. 1-74% 3. 75-94% 4. 95-100% 5. Unknown match rate/have not analyzed
8	For the following types of student-level data, in which direction does data move between systems? Note that some options are not applicable (e.g., transcript data move from the SEA to workforce, not from workforce to the SEA). If you select an option that is not applicable in the online system, you will receive an error message. Select an applicable option in the online system to clear the error message. 1. Demographic 2. Enrollment 3. Transcript data 4. Earnings and wages 5. Employment status 6. Occupation 7. Industry of employment	1. SEA to workforce 2. Workforce to SEA 3. This student-level information is NOT shared
10	For each of the following employment programs or types of employment information, please indicate the range of the match rate when individual students' postsecondary records are matched with their respective workforce records at least annually. 1. Unemployment insurance wage records 2. Unemployment benefits claim data 3. WIA adult or dislocated worker program 4. WIA youth program 5. Adult basic and secondary education 6. Wagner-Peyser Act employment services 7. TANF	1. Do not match (0%) 2. 1-74% 3. 75-94% 4. 95-100% 5. Unknown match rate/have not analyzed

Question number	Question wording	Response options
11	For the following types of student-level data, in which direction does data move between state agencies? 1. Demographic 2. Enrollment 3. Transcript data 4. Financial data 5. Postsecondary degree completion 6. Postsecondary type of degree 7. Earnings and wages 8. Employment status 9. Occupation 10. Industry of employment	1. Postsecondary to workforce 2. Workforce to postsecondary 3.This student-level information is NOT shared

Source: 2013 Data Quality Campaign Survey Instrument. | GAO-15-27

Table 9: DQC Survey Questions We Analyzed to Answer Objective 2

Question number	Question wording	Response options
37	Which of the following types of reports using student-level longitudinal data does the SEA directly or indirectly produce (e.g., the SEA may provide support to districts, regional service centers, institutions of higher education, higher education coordinating bodies, or others to indirectly produce the reports)? Please select all that apply.	1. Diagnostic reports 2. Early warning reports (i.e., reports designed to identify students who are most likely to be at risk of academic failure or dropping out of school) 3. Growth reports 4. College and career readiness reports (i.e., reports designed to identify students who are on track for readiness or success in college or careers) 5. None of the above
40	Which of the following types of reports using aggregate-level longitudinal data does the SEA directly or indirectly produce (e.g., the SEA provides support to districts, regional service centers, institutions of higher education, higher education coordinating bodies, or others to indirectly produce the reports)? Please select all that apply.	1. High school feedback reports. Example: Reports that provide information on how the graduates of a district, school, or program fare in postsecondary, such as the number of graduates from a given high school who enrolled in higher education and the college grades they earned. 2. Internal K–12 feedback reports. Examples: High school to middle school, middle school to elementary school, elementary school to early childhood. 3. Growth reports. Example: Reports that provide analysis of students' academic growth by grade level and subject, including value-added reports. 4. Cohort graduation/completion reports. Example: Reports that provide longitudinal graduation rates disaggregated by students' prior achievement and other suitable at-risk indicators (student attendance, mobility, course patterns, etc.). 5. Predictive reports. Examples: Reports that show differences in students' college enrollment rates and grades based on their test scores, course grades, and graduation plans in high school. 6. College and career readiness reports. Example: School-level report card that includes college and career readiness indicators such as participation and performance in college readiness courses (e.g., AP, IB) or assessments (e.g., SAT, ACT) or the number and percentage of the high school's graduates who attend in-state public colleges within one year of graduation and who need remediation in math or English. 7. None of the above.
43	With which of the following types of organizations has the state developed a purposeful research agenda (i.e., articulated and prioritized research and/or policy questions)?	1. Institutions of higher education or higher education coordinating bodies 2. Independent researchers (not employees of the state) 3. Intermediaries (e.g., AIR, Rand, Consortium on Chicago School Research) 4. None, our state does not have a research agenda developed in conjunction with other organizations

Question number	Question wording	Response options
44	Is there a process by which researchers that are not employees of the state can propose their own studies for approval?	1. Yes 2. No
53	What outreach tools does the state use to communicate the availability of data to noneducator stakeholders (i.e., stakeholders who are not teachers, principals, and superintendents)? Please select all that apply.	1. Traditional outreach (e.g., public service announcements; paid advertisements; press conferences and news releases; general SEA materials including information on data, such as websites and standard school reports) 2. In-person opportunities (e.g., meetings, presentations, conferences) 3. Electronic or social media (e.g., Facebook, Twitter, Plaxo, blogs, email blasts/listserv announcements, webinars) 4. None of the above
66	Is your state using data from its longitudinal data system to support parent and student decisionmaking in any of the following ways? Please select all that apply.	1. To calculate and share information with schools about students' eligibility for state or federal aid or merit-based aid (e.g., state scholars program) 2. To pull or request transcript information for students applying for state or federal aid or merit-based aid 3. To provide parents and students with information about students' progress meeting state college enrollment and placement requirements 4. To provide a data tool for students and parents that supports the combination of state and local data to inform the development of a customized student learning path (e.g., Zangle Services) 5. None of the above
69	How is the state (e.g., SEA, state policymakers) using student growth information to inform continuous improvement or decisionmaking, whether or not this information is publicly reported? Please select all that apply.	1. As part of the state's approved ESEA flexibility waiver for school accountability 2. Inform school turnaround efforts 3. Inform resource allocation 4. Inform policy making 5. Include in teacher or principal evaluation 6. Include in merit pay, differentiated pay, or educator compensation 7. Design or evaluate professional development programs 8. Design or evaluate educator preparation programs (including alternative routes) 9. Evaluate intervention strategies or programs 10. Identify and reward schools that demonstrate high growth 11. None of the above
74	Has the state initiated Skills Gap analyses, using available Labor Market Information (LMI) and industry feedback, to assess alignments between education and workforce programs and labor market demands?	1. Yes 2. No

Question number	Question wording	Response options
79	Does the state regularly report aggregate data on credential, employment, and transfer outcomes across all of the state's education (e.g., adult basic education, career and technical education) and workforce programs to the Governor and state legislature?	1. Yes 2. No
134	What publicly available statistics or indicators are created based on data from the K–12 and postsecondary linkages? Please select all that apply.	1. Postsecondary enrollment rates 2. Postsecondary remediation rates 3. None of the above
138	Can the state use the K–12 and postsecondary linked data in any of the following ways? Please select all that apply.	1. Create customized reports (e.g., in response to researcher request) 2. Create tailored reports for different stakeholder groups (e.g., educators, parents, students) 3. None of the above
139	Does the state use the linked K–12 and postsecondary data for purposes beyond compliance reporting?	1. Yes 2. No

Source: 2013 Data Quality Campaign Survey Instrument. | GAO-15-27

Appendix III: Statewide Longitudinal Data Systems (SLDS) and Workforce Data Quality Initiative (WDQI) Grantees

Table 10: Total Number of SLDS or WDQI Grants and Amount Awarded, Fiscal Years 2006-2013

Grantee	Total Number of SLDS Grants Awarded	Total Number of WDQI Grants Awarded	Total SLDS & WDQI Grant Funds Awarded (in dollars)
Alaska	2	0	$7,506,757
Arizona	2	0	10,921,224
Arkansas	3	1	19,057,980
California	2	0	9,255,445
Colorado	2	0	21,653,636
Connecticut	2	1	5,261,921
DC	2	0	9,738,500
Delaware	1	0	4,616,250
Florida	3	1	15,002,890
Georgia	1	0	8,942,640
Hawaii	2	1	7,862,946
Idaho	2	1	10,018,152
Illinois	2	1	21,869,776
Indiana	2	1	10,243,333
Iowa	2	1	13,524,740
Kansas	3	1	17,963,148
Kentucky	3	1	13,200,861
Louisiana	1	1	5,056,373
Maine	2	1	11,542,231
Maryland	3	1	16,644,377
Massachusetts	2	1	19,966,194
Michigan	3	1	20,142,192
Minnesota	2	1	16,684,225
Mississippi	2	1	11,924,999
Missouri	1	1	9,857,686
Montana	2	0	9,776,318
Nebraska	2	2	9,895,861
Nevada	2	0	9,999,965
New Hampshire	2	0	8,165,663
New Jersey	1	2	6,075,704
New York	2	0	27,515,288
North Carolina	2	1	10,795,891

Grantee	Total Number of SLDS Grants Awarded	Total Number of WDQI Grants Awarded	Total SLDS & WDQI Grant Funds Awarded (in dollars)
North Dakota	2	1	11,666,988
Ohio	3	2	15,803,007
Oklahoma	1	1	5,997,082
Oregon	3	1	20,033,822
Pennsylvania	3	1	25,395,896
Puerto Rico	1	0	4,665,708
Rhode Island	2	1	9,667,933
South Carolina	2	1	20,975,280
South Dakota	1	1	3,980,692
Tennessee	1	0	3,226,313
Texas	2	1	27,071,876
U.S. Virgin Islands	1	0	2,606,688
Utah	2	0	14,179,499
Vermont	1	0	4,947,261
Virginia	2	2	25,747,280
Washington	2	1	24,283,758
West Virginia	1	0	4,798,697
Wisconsin	3	0	22,442,310
Wyoming	0	1	722,717

Source: GAO analysis of Department of Education data. | GAO-15-27

Notes: SLDS grantees received awards between fiscal years 2006 and 2012. The Department of Education did not award grants for fiscal years 2008, 2010, 2011, or 2013. WDQI grantees received awards between fiscal years 2010 and 2013.

Appendix IV: Specific Programs Matched by Statewide Longitudinal Data Systems (SLDS) and Workforce Data Quality Initiative (WDQI) Grantees between Sectors

Table 11: Specific Programs Matched by SLDS or WDQI Grantees between the K-12 and Early Education Sectors, Any Match Rate, 2013

Grantee	Subsidized Child Care	Early Intervention	Special Education	State Prekindergarten	Head Start/Early Head Start
Alaska	-	-	x	x	x
Arizona	-	-	-	-	-
Arkansas	x	x	x	x	x
Colorado	-	-	x	x	-
Connecticut	x	x	x	x	x
DC	x	x	x	x	x
Delaware	-	x	x	x	x
Florida	-	x	x	x	-
Georgia	x	x	x	x	x
Hawaii	-	x	x	-	-
Idaho	-	x	x	x	-
Illinois	-	x	x	x	-
Indiana	-	x	x	x	-
Iowa	-	x	-	x	-
Kansas	-	x	x	x	x
Kentucky	x	x	x	x	x
Louisiana	-	-	x	x	-
Maine	-	-	x	x	x
Maryland	x	x	x	x	x
Massachusetts	x	-	x	x	x
Michigan	-	x	x	x	-
Minnesota	-	x	x	-	-
Mississippi	-	-	-	x	x
Missouri	x	x	x	x	x
Montana	-	x	x	x	-
Nebraska	-	x	x	x	x
Nevada	-	-	x	x	-
New Hampshire	-	x	x	x	-
New Jersey	-	x	x	x	x
New York	-	x	x	x	-
North Carolina	-	-	x	x	x
North Dakota	-	-	x	x	x
Ohio	-	x	x	x	x

Grantee	Subsidized Child Care	Early Intervention	Special Education	State Prekindergarten	Head Start/Early Head Start
Oklahoma	-	-	x	x	x
Oregon	-	x	x	x	x
Pennsylvania	-	x	-	x	-
Rhode Island	x	x	x	x	x
South Carolina	-	-	-	-	-
South Dakota	-	-	-	-	-
Tennessee	-	-	x	x	-
Texas	-	-	x	x	-
Utah	-	x	x	-	-
Vermont	-	-	-	-	-
Virginia	x	-	x	x	x
Washington	-	x	x	x	-
West Virginia	-	-	-	-	-
Wisconsin	-	-	x	x	-
Wyoming	-	-	x	x	-

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Table 12: Specific Types of Postsecondary Institutions Matched to K-12 Data by SLDS or WDQI Grantees, Any Match Rate, 2013

Grantee	Less Than 2-Year Public	Less Than 2-Year Private Not-for-profit	Less Than 2-Year Private For-profit	2-Year Public	2-Year Private Not-for-profit	2-Year Private For-profit	4-Year and Above Public	4-Year and Above Private Not-for-profit	4-Year and Above Private For-profit
Alaska	-	-	-	x	-	-	x	-	x
Arizona	-	-	-	-	-	-	-	-	-
Arkansas	x	x	-	x	x	-	x	x	-
Colorado	-	-	-	x	-	-	x	-	-
Connecticut	-	-	-	-	-	-	-	-	-
DC	x	x	x	x	x	x	x	x	x
Delaware	-	-	-	x	x	x	x	x	x
Florida	x	x	-	x	x	-	x	x	-
Georgia	x	-	-	x	-	-	x	-	-
Hawaii	x	x	x	x	x	x	x	x	x
Idaho	x	-	-	x	-	-	x	-	x
Illinois	-	-	-	x	-	-	x	-	-
Indiana	x	x	x	x	x	x	x	x	x
Iowa	-	-	-	x	-	-	x	-	-
Kansas	x	x	x	x	x	x	x	x	x
Kentucky	-	-	-	x	-	-	x	x	-
Louisiana	x	x	x	x	x	x	x	x	x
Maine	-	-	-	x	x	x	x	x	x
Maryland	x	x	x	x	x	x	x	x	x
Massachusetts	x	x	x	x	x	x	x	x	x
Michigan	-	-	-	x	x	x	x	x	x
Minnesota	x	x	x	x	x	x	x	x	x
Mississippi	x	-	-	x	-	-	x	-	-
Missouri	x	-	-	x	-	-	x	-	-
Montana	-	-	-	x	-	-	x	-	-
Nebraska	-	-	-	-	-	-	-	-	-
Nevada	-	-	-	x	-	-	x	-	-
New Hampshire	-	-	-	x	-	-	x	-	-
New Jersey	x	x	x	x	x	x	-	x	x
New York	x	x	x	x	x	x	x	x	x
North Carolina	x	-	-	x	x	-	x	x	-
North Dakota	-	-	-	x	-	-	x	-	-

Grantee	Less Than 2-Year Public	Less Than 2-Year Private Not-for-profit	Less Than 2-Year Private For-profit	2-Year Public	2-Year Private Not-for-profit	2-Year Private For-profit	4-Year and Above Public	4-Year and Above Private Not-for-profit	4-Year and Above Private For-profit
Ohio	-	-	-	-	-	-	-	-	-
Oklahoma	x	-	-	x	-	-	x	-	-
Oregon	-	-	-	x	x	x	x	x	x
Pennsylvania	x	-	-	x	-	-	x	-	-
Rhode Island	-	-	-	x	x	x	x	x	x
South Carolina	-	-	-	x	-	-	x	x	-
South Dakota	-	-	-	x	-	-	x	-	-
Tennessee	x	x	-	x	x	-	x	x	-
Texas	-	-	-	x	x	x	x	x	x
Utah	x	-	-	x	-	-	x	-	-
Vermont	-	-	-	-	-	-	-	-	-
Virginia	x	x	x	x	x	x	x	x	x
Washington	x	-	-	x	-	-	x	-	-
West Virginia	-	-	-	x	-	-	x	-	-
Wisconsin	x	-	-	x	-	-	x	-	-
Wyoming	-	-	-	x	-	-	x	-	-

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Table 13: Specific Programs Matched by SLDS or WDQI Grantees between the K-12 and Workforce Sectors, Any Match Rate, 2013

Grantee	Unemployment Insurance Wage Records	Unemployment Benefits Claims Data	WIA Adult or Dislocated Worker Program	WIA Youth Program	Adult Basic and Secondary Education	Wagner-Peyser Act Employment Services	TANF
Alaska	x	x	x	x	x	-	-
Arizona	-	-	-	-	-	-	-
Arkansas	x	x	x	x	x	x	x
Colorado	x	-	-	-	-	-	-
Connecticut	-	-	-	-	x	-	-
DC	x	x	x	x	x	-	x
Delaware	x	-	-	-	x	-	-
Florida	x	x	x	x	x	x	x
Georgia	-	-	-	-	-	-	-
Hawaii	-	-	-	-	-	-	-
Idaho	-	-	-	-	-	-	-
Illinois	-	-	-	-	-	-	-
Indiana	x	x	-	-	x	-	x
Iowa	x	-	-	-	-	-	-
Kansas	-	-	-	-	-	-	x
Kentucky	x	-	-	-	x	-	-
Louisiana	-	-	x	x	-	-	-
Maine	x	-	-	-	x	-	-
Maryland	x	-	-	-	x	-	-
Massachusetts	-	-	-	-	x	-	-
Michigan	-	-	-	-	x	-	x
Minnesota	-	-	-	-	x	-	-
Mississippi	x	x	x	x	x	x	x
Missouri	x	x	x	x	x	-	-
Montana	-	-	-	-	x	-	-
Nebraska	-	-	-	-	-	-	-
Nevada	-	-	-	-	x	-	-
New Hampshire	-	-	-	-	-	-	-
New Jersey	-	-	-	-	-	-	-
New York	-	-	-	-	-	-	-
North Carolina	x	x	x	x	x	x	x
North Dakota	x	-	-	-	x	-	-

Grantee	Unemployment Insurance Wage Records	Unemployment Benefits Claims Data	WIA Adult or Dislocated Worker Program	WIA Youth Program	Adult Basic and Secondary Education	Wagner-Peyser Act Employment Services	TANF
Ohio	-	-	-	-	-	-	-
Oklahoma	-	-	-	-	-	-	x
Oregon	x	x	x	x	x	x	x
Pennsylvania	-	-	-	-	-	-	-
Rhode Island	x	-	-	-	x	-	x
South Carolina	-	-	-	-	-	-	x
South Dakota	-	-	-	-	-	-	-
Tennessee	x	x	x	x	x	x	-
Texas	x	-	-	-	x	-	-
Utah	x	-	-	-	x	-	-
Vermont	-	-	-	-	-	-	-
Virginia	-	-	-	-	-	-	-
Washington	x	x	x	x	x	x	x
West Virginia	-	-	-	-	-	-	-
Wisconsin	-	-	-	-	-	-	-
Wyoming	-	-	-	-	-	-	-

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: WIA is the Workforce Investment Act of 1998. TANF is the Temporary Assistance for Needy Families program.

Table 14: Specific Programs Matched by SLDS or WDQI Grantees between the Postsecondary and Workforce Sectors, Any Match Rate, 2013

Grantee	Unemployment Insurance Wage Records	Unemployment Benefits Claims Data	WIA Adult or Dislocated Worker Program	WIA Youth Program	Adult Basic and Secondary Education	Wagner Peyser Act Employment Services	TANF
Alaska	x	x	x	x	x	-	-
Arizona	-	-	-	-	-	-	-
Arkansas	x	x	x	x	x	x	x
Colorado	x	-	x	x	-	x	-
Connecticut	x	-	-	-	-	-	-
DC	-	-	-	-	-	-	x
Delaware	-	-	-	-	-	-	-
Florida	x	x	x	x	x	x	x
Georgia	-	-	-	-	-	-	-
Hawaii	x	-	-	-	-	-	-
Idaho	-	-	-	-	x	-	-
Illinois	x	-	x	-	x	-	-
Indiana	x	x	-	-	x	-	-
Iowa	x	-	-	-	-	-	-
Kansas	x	-	-	-	x	-	-
Kentucky	x	-	-	-	x	-	-
Louisiana	-	-	-	-	-	-	-
Maine	x	-	-	-	x	-	-
Maryland	-	-	-	-	-	-	-
Massachusetts	x	-	-	-	x	-	-
Michigan	-	-	-	-	x	-	-
Minnesota	x	-	x	x	x	-	-
Mississippi	x	x	x	x	x	x	x
Missouri	-	-	x	x	x	x	-
Montana	x	-	-	-	-	-	-
Nebraska	x	-	-	-	-	-	-
Nevada	-	-	-	-	-	-	-
New Hampshire	-	-	-	-	-	-	-
New Jersey	-	-	-	-	-	-	-
New York	-	-	-	-	-	-	-
North Carolina	x	x	x	x	x	x	x
North Dakota	x	-	-	-	x	-	-

Grantee	Unemployment Insurance Wage Records	Unemployment Benefits Claims Data	WIA Adult or Dislocated Worker Program	WIA Youth Program	Adult Basic and Secondary Education	Wagner Peyser Act Employment Services	TANF
Ohio	x	-	-	-	x	-	-
Oklahoma	x	-	-	-	-	-	-
Oregon	x	x	x	x	x	x	x
Pennsylvania	-	-	-	-	-	-	-
Rhode Island	x	-	-	-	x	-	x
South Carolina	-	-	-	-	-	-	-
South Dakota	x	-	-	-	-	-	-
Tennessee	x	x	x	x	x	x	-
Texas	x	-	x	x	x	x	x
Utah	x	-	x	x	x	x	x
Vermont	-	-	-	-	-	-	-
Virginia	x	-	-	-	x	-	-
Washington	x	x	x	x	x	x	x
West Virginia	x	-	x	-	x	-	-
Wisconsin	x	-	-	-	x	-	-
Wyoming	-	-	x	x	x	-	-

Source: GAO analysis of 2013 Data Quality Campaign (DQC) survey data. | GAO-15-27

Notes: WIA is the Workforce Investment Act of 1998. TANF is the Temporary Assistance for Needy Families program.

Appendix V: GAO Contact and Staff Acknowledgments

GAO Contact

Jacqueline M. Nowicki, Director, 617-788-0580 or nowickij@gao.gov

Staff Acknowledgments

In addition to the contact named above, Janet Mascia, Assistant Director, Jennifer Gregory, and Nisha R. Hazra made key contributions to this report. Also contributing to this report were Deborah Bland, David Chrisinger, Alex Galuten, Amanda Miller, Jeffrey G. Miller, Mimi Nguyen, Yunsian Tai, and Walter Vance.

Glossary

This glossary is provided for reader convenience. It is not intended as a definitive, comprehensive glossary of related terms.

Aggregate level data	Group statistics (numbers, percentages, averages, etc.) based on individual student data.
College and career readiness reports	Reports designed to identify students who are on track for readiness or success in college or careers.
Data governance	The exercise of decision-making and authority for data-related matters using agreed-upon rules that describe who can take what actions with what information and when, under what circumstances, and using what methods.
Diagnostic report	Information on individuals designed to identify each student's strengths and academic needs.
Early education sector	Programs that serve children prior to kindergarten. Programs include: early intervention, Head Start/Early Head Start, state prekindergarten, special education, and subsidized child care.
Early warning report	A report designed to identify students who are most likely to be at risk of academic failure or dropping out of school.
Feedback report	Information on outcomes for students after they graduate from a school or district.
Growth report	A report that shows changes in the achievement of the same students over time.
K-12 sector	Elementary and secondary education.
Postsecondary education sector	Institutions of higher education. Types of institutions include: less than 2-year public, 2-year public, 4-year and above public, less than 2-year private not-for-profit, 2-year private not-for-profit, 4-year and above private not-for-profit, less than 2-year private for-profit, 2-year private for-profit, and 4-year and above private for-profit.
Predictive report	A report that shows how students' success later in the education/workforce pipeline is related to the status of the same students earlier in the pipeline.
Matching	Reliably connecting the same individual record in two or more databases.

Match rate	The percent of unique individual records reliably connected across databases.
Sharing	Exchanging data between two databases, in either direction. Data elements that could be shared between early education and K-12 include: demographic, family characteristics, program participation, child-level development data; between K-12 and postsecondary: demographic, college readiness assessment scores, college placement assessment scores, high school transcript data, postsecondary enrollment, postsecondary remediation status, postsecondary progress, postsecondary credits earned, postsecondary enrollment intensity, postsecondary outcomes; between K-12 and workforce: demographic, enrollment, transcript data, earnings and wages, employment status, occupation, industry of employment; between post-secondary and workforce: demographic, enrollment, transcript data, financial aid, postsecondary degree completion, earnings and wages, employment status, occupation, industry of employment.
Workforce sector	Programs that serve individuals in the workforce. Programs include: adult basic and secondary education, TANF, unemployment benefits claims data, unemployment insurance wage records, Wagner-Peyser Act employment, WIA adult or dislocated workers program, and WIA youth program.